Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **410,** 2653–2661 (2011) doi:10.1111/j.1365-2966.2010.17641.x

# Accretion shocks and cold filaments in galaxy formation

Andrew J. Benson[1]★ and Richard Bower[2]

[1]*Mail Code 350-17, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA*
[2]*Institute for Computational Cosmology, University of Durham, Durham DH1 3LE*

Accepted 2010 September 3. Received 2010 August 26; in original form 2010 April 7

## ABSTRACT

A generic expectation for gas accreted by high-mass haloes is that it is shock-heated to the virial temperature of the halo. In low-mass haloes, or at high redshift, however, the gas cooling rate is sufficiently rapid that an accretion shock is unlikely to form close to the virial radius. Instead, the accretion shock will form at smaller radii, perhaps close to the central galaxy. Semi-analytic models have always made a clear distinction between the regimes in which accretion is limited by the cooling time of hot gas and by the dynamical time-scale of the halo, using simple estimates of the mass-scale at which the transition from rapid to slow cooling occurs. In this work, we revisit this issue using the latest understanding and calibration of accretion shock formation. Starting from the well-established GALFORM code, we investigate the effect of accounting for the presence or otherwise of an accretion shock close to the virial radius using the shock stability model of Birnboim & Dekel. As expected, when we modify the code so that there is no effective feedback from galaxy formation, we find that so-called 'cold-mode' accretion is the dominant channel for feeding gas into the galaxies at high redshifts, such that 90 per cent of baryons in galaxies (averaged over all galaxies) arrive via this channel. The mass-scale at which the rapid to slow cooling transition occurs is significantly affected at high redshifts and accretion rates become dominated by cold-mode accretion. However, the impact of this change in the cooling channel on galaxies properties is mitigated by compensating effects in the star formation and feedback cycle. When effective feedback, which reheats gas from galaxies to the virial temperature but which allows no gas to escape from a halo, is included in the model, we find that the 'cold mode' is even less apparent because of the presence of gas ejected from the galaxy's disc, although it can still contribute almost 50 per cent of the net inflow rate when averaged over all galaxies. Thus, the inclusion of the latest calibration of accretion shock physics makes little difference to basic results from earlier semi-analytic models, which used a simpler treatment. We conclude that this 'cold-mode' physics is already adequately accounted for in semi-analytic models and that feedback represents a much larger uncertainty than any of these effects.

**Key words:** galaxies: evolution – galaxies: formation – galaxies: general.

## 1 INTRODUCTION

The process of galaxy formation must begin with gas, initially distributed rather smoothly, collapsing to high densities. Furthermore, to sustain ongoing star formation in galaxies requires a continued input of gas over their lifetimes. As such, the question of how galaxies get their gas has received a great deal of attention over the history of galaxy formation studies. While the initial stages of this collapse are purely gravitational (the gas being dragged along by the gravitationally dominant dark matter), after halo formation, hydrodynamic forces come into play and further collapse is mitigated by the interplay of gravity, hydrodynamics and cooling processes.

An accretion shock is a generic expectation whenever the gas accretes supersonically as it will do if the halo virial temperature exceeds the temperature of the accreting gas (Binney 1977). Models of accretion shocks have been presented by several authors (Bertschinger 1985; Tozzi & Norman 2001; Voit et al. 2003; Book & Benson 2010) with the general conclusion that the shock occurs at a radius comparable to (or perhaps slightly larger than) the virial radius when cooling times are long compared to dynamical times. In the other limit of short cooling times, it has long been understood that the shock must instead form at much smaller radii, close to the forming galaxy (Rees & Ostriker 1977; White & Frenk 1991).

★E-mail: abenson@caltech.edu

For example, according to Rees & Ostriker (1977): 'Unless pregalactic clouds collapse in an exceedingly homogeneous fashion, their kinetic energy of infall will be thermalized by shocks before collapse has proceeded by more than a factor $\sim$2. What happens next depends on the relative value of the cooling and collapse timescales. Masses in the range $10^{10}$–$10^{12}\,M_\odot$ cool so efficiently that they always collapse at the free-fall rate,[1] and probably quickly fragment into stars. Larger masses may, however, experience a quasistatic contraction phase . . .'. Thus, Rees & Ostriker (1977) clearly understood the difference between the rapid inflow and slow cooling regimes, and correctly identified the transition mass, suggesting that this be identified with the characteristic stellar mass of galaxies. Accretion in these two regimes may be expected to result in very different spatial and spectral distributions of cooling radiation, leading to the possibility of observationally distinguishing the two types of accretion (Fardal et al. 2001).

The distinction between these two regimes has always been an integral part of (semi-)analytic models of galaxy formation, beginning with Rees & Ostriker (1977). For example, White & Frenk (1991) introduced a transition between rapid- and slow-cooling regimes at the point where cooling and virial radii become equal, or, equivalently, the point at which cooling and dynamical times at the halo virial radius become equal. In the rapid-cooling regime, the accretion rate of gas into the central galaxy was then determined by the cosmological infall rate, while in the slow cooling regime, the accretion rate was determined by the cooling time in the gas. Their fig. 2 illustrates that the rapid-cooling regime will occur in low-mass haloes and at high redshifts. All subsequent semi-analytic models of galaxy formation (e.g. Kauffmann, White & Guiderdoni 1993; Cole et al. 1994) have adopted this prescription, or some variant of it, and it has also been validated by one-dimensional hydrodynamical simulations (e.g Forcada-Miro & White 1997). The validity of this prescription has been confirmed by studies, which compared its predictions for the condensed masses of galaxies with those from smoothed particle hydrodynamics simulations across the boundary of the rapid to slow cooling transition (Benson et al. 2001; Yoshida et al. 2002; Helly et al. 2003), although it should be noted that the accuracy of these comparisons is less than that at which semi-analytic models are now being used.

Over the past decade, further insight into the nature of gas accretion and cooling has arisen, largely as a result of numerical hydrodynamical simulations. Cosmological simulations (Katz et al. 2003; Kereš et al. 2005; see also Katz & Gunn 1991) demonstrated directly that accretion via cold filaments is the dominant channel of gas supply to galaxies over a wide range of masses. Furthermore, simple models and one-dimensional hydrodynamical simulations (Birnboim & Dekel 2003) built upon the earlier work on accretion shocks and demonstrated explicitly that shock stability is governed by the properties of the currently accreting gas. This improved upon the simple time-scale arguments made previously and permitted the process of gas collapse to be followed into the vicinity of the galaxy itself when cooling time-scales were short. Later generations of three-dimensional simulations demonstrated that the reduced cooling times in very high density contrast accreting filaments could modify the shock stability criteria and effectively prevent shock propagation in such dense regions (Kereš et al. 2005; Dekel et al. 2009). As a result, the transition between rapid and slow regimes becomes less distinct, allowing cold filaments to continue to supply gas to galaxies even in haloes with a dominant hot atmospheres. As we will discuss below, the channel by which a galaxy receives its gas may be important even if the accretion rate is unchanged. The delivery mechanism may affect the rate of angular momentum accretion on to a galaxy and will affect the detectability of the accreting gas (e.g. Benson et al. 2000; Fardal et al. 2001; Crain et al. 2010; Goerdt et al. 2010).

Recent work has once again focused on the formation of accretion shocks. Recent three-dimensional hydrodynamical simulations (Fardal et al. 2001; see also Kereš et al. 2005; Ocvirk, Pichon & Teyssier 2008; Kereš et al. 2009) have suggested that a significant fraction of gas in low-mass galaxies has never been shock-heated (at least within regions adequately resolved by the simulations; for example, numerical simulations may not adequately resolve shocks in the radiative regime due to artificial viscosity, numerical diffusion and other numerical artefacts; Agertz et al. 2007; Creasy et al. in preparation). Motivated by these results Birnboim & Dekel (2003) developed an analytic treatment of accretion shock stability.[2] The accretion shock relies on the presence of a stable atmosphere of post-shock gas to support itself. If cooling times in the post-shock gas are sufficiently short, this atmosphere cools and collapses and can no longer support the shock. The shock therefore shrinks to smaller radii, where it can be stable. For cosmological haloes, this implies that shocks can only form close to the virial radius in haloes with mass greater than $10^{11}\,M_\odot$ for primordial gas (or around $10^{12}\,M_\odot$ for gas of solar metallicity). These values are found to depend only weakly on redshift and are in good agreement with the results of hydrodynamical simulations. It is crucial to note that these new criteria are equivalent to that of White & Frenk (1991) up to factors of the order of unity, with one small caveat, as we will discuss in more detail below.

As a result, in low-mass haloes gas tends to accrete into haloes 'cold' – never being shock-heated to the virial temperature and proceeding to flow along filaments towards the centre of the halo where it will eventually shock.[3] Haloes, which do support shocks close to the virial radius, are expected to contain a quasi-hydrostatic atmosphere of hot gas. The structure of this atmosphere is determined by the entropy that the gas gains at the accretion shock and that may be later modified by radiative cooling and feedback (Voit et al. 2003; McCarthy et al. 2007). In practice, the transition from rapid to slow cooling regimes is not sharp – haloes able to support a shock at their virial radius still contain some unshocked gas, because the haloes retain a memory of past accretion and because cold filaments may penetrate through the hot halo. At high redshifts in particular, the 'cold' accretion mode may be active even in haloes whose accretion of gas is primarily via an accretion shock close to the virial radius (Kereš et al. 2009).

The consequences of rapid versus slow cooling regimes for the properties of the galaxy forming warrant further study. As Croton et al. (2006) have stressed, the absence of a more detailed treatment of the rapid-cooling regime may not be important, since, by definition, the gas accretion rate in small haloes is limited by the growth of the halo rather than by the system's cooling time. In contrast, Brooks et al. (2009) demonstrate in hydrodynamical simulations that in the rapid-cooling regime, accreted gas can reach the galaxy more rapidly, by virtue of the fact that it does not have to cool but instead merely has to free fall to the centre of the halo (starting with

[1] We will argue below that they in fact collapse somewhat faster than this.

[2] Unpublished work by Forcada-Miro & White (1997), also utilizing a one-dimensional hydrodynamics code, reached similar conclusions.

[3] While this picture seems reasonable on theoretical grounds, it as yet has little direct observational support (Steidel et al. 2010).

a velocity comparable to the virial velocity). This results in earlier star formation than if all gas were assumed to be initially shock-heated to the virial temperature close to the virial radius of the halo. It is also clear that the situation needs to be carefully reassessed in the presence of effective feedback schemes that prevent excessive star formation, particularly in the high-redshift universe.

In this work, we implement a treatment of rapid/slow cooling regimes into the GALFORM semi-analytic model of galaxy formation, following the methodology of Birnboim & Dekel (2003). We emphasize that GALFORM has always made the distinction between rapid- and slow-cooling regimes. We are simply adopting an updated calibration of the transition mass between these regimes and accounting for the shorted time-scale on which gas may reach the central galaxy in the rapid-cooling regime. The two schemes are not exactly equivalent, however, as we discuss in Section 2. This will allow us to assess the importance of the accretion shocks for cosmological populations of galaxies across a range of redshifts and to suggest additional studies of this regime, which might improve our understanding of its role in the process of galaxy formation. An important aspect is that we are able to explore how feedback and reheating of galactic gas moderate the hot and cold gas content of haloes. We note that Cattaneo et al. (2006) have previously explored a simpler implementation of accretion shocks in a semi-analytic model of galaxy formation. They parametrized the role of accretion shocks by defining a critical mass-scale (which had some dependence on redshift) above which cooling switched from 'cold accretion' to slow-cooling regimes. While motivated by the results of the Birnboim & Dekel (2003) calculation, this parametrization did not capture the full generality of that work as we will attempt to do herein.

The remainder of this paper is arranged as follows. In Section 2, we describe our implementation of the rapid-cooling regime in the GALFORM semi-analytic model. In Section 3, we present our results and, finally, in Section 4, we discuss their implications.

## 2 RAPID-COOLING REGIME MODEL

We implement the rapid-cooling regime in our semi-analytic model using the results of Birnboim & Dekel (2003). Birnboim & Dekel (2003) derive a simple criterion for accretion shock stability, which states that

$$\epsilon_{\rm s} < \epsilon_{\rm s,crit}, \quad \text{where} \quad \epsilon_{\rm s} = \rho_0 r_{\rm s} \Lambda(T_1)/u_0^3, \tag{1}$$

where $\rho_0$ is the pre-shock gas density, $r_{\rm s}$ is the shock radius, $\Lambda(T)$ is the usual cooling function,[4] $u_0$ is the pre-shock inflow velocity of the gas and $T_1 = (3/16)\mu m_{\rm H} u_0^2/k_{\rm B}$ is the post-shock gas temperature. The left-hand side of this expression is equivalent, to an order of magnitude, to the ratio of sound-crossing and cooling times in the post-shock gas. This provides some physical insight into this condition: if the post-shock gas can cool too quickly, sound waves cannot communicate across the halo and thereby form a hydrostatic atmosphere, which can support a shock front). Consequently, this new criterion is identical to that proposed by White & Frenk (1991) up to a factor of the order of unity, with one small caveat. The White & Frenk (1991) condition, as used in GALFORM and other semi-analytic models, depends explicitly on the gas content of the dark matter halo, since it utilizes the cooling radius, which itself depends on the density of gas in the halo. As such, the White & Frenk (1991) condition will depend on the accretion and cooling history of a given halo and, possibly, on whether any gas is returned to the halo from the galaxy via feedback processes. The Birnboim & Dekel (2003) criterion, as expressed above, depends only on the properties of the currently accreting gas. While this distinction may be a minor one in practice, scenarios where it is important can be constructed. For example, consider a halo that has cooled much of its gas and so is filled with a low-density atmosphere in which the cooling radius is smaller than the virial radius. The White & Frenk (1991) criterion then places this halo in the slow-cooling regime but, if the halo is of sufficiently low mass, the Birnboim & Dekel (2003) criterion may place it in the rapid-cooling regime. In such a case, we can envision filaments of cold material penetrating the diffuse hot atmosphere as has been seen in simulations (e.g. Dekel et al. 2009). Whether or not this lack of dependence on the current hot halo properties in the Birnboim & Dekel (2003) method (which assumes spherical symmetry) is physically correct, it is different from our previous method and so should be kept in mind when examining model results. The stability parameter $\epsilon_{\rm s}$ was found to be[5] 0.0126 by Birnboim & Dekel (2003) for a gas with adiabatic index $\gamma = 5/3$. Clearly this value is somewhat uncertain, given the approximations made, but Birnboim & Dekel (2003) justify it on the basis of hydrodynamical simulations. At each time-step of our calculation, we compute this criterion for each dark matter halo using the pre-accretion model described by Birnboim & Dekel (2003) in their section 5.1. We evaluate this expression at the virial radius, that is, $r_{\rm s} = r_{\rm vir}$, with an infall velocity $u_0^2 = {\rm G}M_{\rm vir}/r_{\rm vir}$ and a pre-shock density $\rho_0$ computed following equation (48) of Birnboim & Dekel (2003). We assume that a fraction, $f_{\rm slow}$, of any gas accreted from the intergalactic medium (IGM) during that time-step is added to the usual hot gas reservoir of the halo (which is assumed to be heated to the virial temperature by the virial shock). We adopt a simple model in which

$$f_{\rm slow} = [1 + \exp(\{\epsilon_{\rm s,crit} - \epsilon_{\rm s}\}/\Delta\epsilon)]^{-1} \tag{2}$$

such that $f_{\rm slow} \rightarrow 1$ for $\epsilon_{\rm s} \ll \epsilon_{\rm s,crit}$ and $f_{\rm slow} \rightarrow 0$ for $\epsilon_{\rm s} \gg \epsilon_{\rm s,crit}$ with the width of the transition being controlled by the parameter $\Delta\epsilon$. There is no physical motivation for this particular functional form – it simply allows for a smooth transition with an adjustable transition width. Cooling and infall of this hot-mode gas then follows the standard treatment of the Benson & Bower (2010) semi-analytic model. The remaining fraction $f_{\rm rapid} (\equiv 1 - f_{\rm hot})$ is instead added to a new cold reservoir in the halo. Gas in this rapid-cooling reservoir is allowed to infall on to the central galaxy of the halo at a rate

$$\dot{M}_{\rm infall,rapid} = \Gamma_{\rm rapid} M_{\rm rapid}/t_{\rm ff}, \tag{3}$$

where $M_{\rm rapid}$ is the current mass of gas in the cold reservoir, $t_{\rm ff}$ is the free-fall time from the halo virial radius to the halo centre and $\Gamma_{\rm rapid}$ is a parameter of the order of unity, which we will use to adjust the infall rate. Specifically, the free-fall time is defined as

$$t_{\rm ff} = \int_0^{r_{\rm vir}} \frac{{\rm d}r}{\sqrt{2[\Phi(r_{\rm vir}) - \Phi(r)]}}, \tag{4}$$

[4] In models with no supernova (SN) driven outflows, we adopt a primordial composition when computing the cooling function and assume collisional ionization equilibrium. This is consistent with the cooling function used in the hydrodynamical simulations with which we compare. When SN-driven outflows are included, we self-consistently compute the enrichment of gas and use a cooling function computed for the appropriate metallicity.

[5] We do not explicitly include the change in the criterion as a function of redshift haloes found by Dekel & Birnboim (2006). This may lead to some overestimate of mass in the rapid mode in low-redshift haloes.

where $\Phi(r)$ is the gravitational potential at radius $r$ in the halo. For Navarro–Frenk–White haloes, $t_{\rm ff} \approx 1.2 r_{\rm vir}/V_{\rm vir}$, where $r_{\rm vir}$ and $V_{\rm vir}$ are the virial radius and velocity of the halo, respectively. Since cold flow gas arrives at the virial radius moving with speed comparable to the virial velocity, we may reasonably expect $\Gamma_{\rm rapid} > 1$. In fact, calculations suggest that cold gas reaches the centre of the halo in around one-half of the free-fall time (Yuval Birnboim, private communication). Since equation (3) results in an exponential decline in the rapid infall reservoir, whereas in reality all of this gas should flow into the galaxy in finite time, we may expect $\Gamma$ to be larger still. The metal and angular momentum content of the cold reservoir are also tracked, and we assume that metals and angular momentum are transferred from the cold reservoir to the central galaxy at rates proportional to the mass infall rate, assuming that the metallicities and specific angular momenta of cold reservoir gas elements are all equal. The specific angular momentum of gas flowing into the halo is assumed to be the same, whether or not it experiences an accretion shock close to the virial radius. This is an assumption and one, which would be useful to test against hydrodynamical simulations. This does not mean that the angular momentum content of galaxies will be independent of whether they are fueled by cooling from a hydrostatic halo or from cold filaments. When accreting from a hydrostatic halo, we follow the usual semi-analytic approach in which cooling occurs from the centre out, resulting in the lowest angular momentum gas cooling first. When being fueled by cold filaments, we simply assume that the gas obtained by a galaxy has the mean specific angular momentum of the cold reservoir. This filamentary fueling should deliver angular momentum to a galaxy earlier than fueling from a hydrostatic halo. We will return to this point later.

If a halo merges with a larger halo, its hot- and cold-mode reservoirs are assigned to the hot and cold reservoirs of the larger halo in just the same way, that is, the gas from the infalling halo is treated as if it were accreted from the IGM. (In practice, this process happens gradually as ram pressure forces remove gas from the satellite halo. However, in most cases, the transfer of gas from a satellite to host happens rapidly after accretion of the satellite.) Our model includes outflows of gas from galaxies, driven by SN explosions. Any such reheated gas is always added to the hot-mode reservoir (after a delay of the order of a halo dynamical time – see Benson & Bower 2010). We assume that this feedback does not affect the content or infall rate of the cold gas reservoir. The rest of our semi-analytic model remains unchanged and follows the methods described by Benson & Bower (2010). We briefly review a few key aspects of the model below.

We generate dark matter merger trees using the algorithm of Parkinson, Cole & Helly (2008), which reproduces the progenitor mass functions (and therefore accretion rates) of haloes in $N$-body simulations. We simulate trees spanning a range of masses from $10^{10}$ to $10^{15}\,h^{-1}\,{\rm M}_{\odot}$, resolving branches down to $5\times10^{9}\,h^{-1}\,{\rm M}_{\odot}$, which we find is sufficient to give converged answers for the quantities studied in this work and which is comparable to the resolution of the simulations to which we compare [e.g. a 20-particle halo in the simulation of Kereš et al. (2009) would have a mass of $7.5\times10^{9}\,h^{-1}\,{\rm M}_{\odot}$].

Following Benson & Bower (2010), rates of cooling from the hot, hydrostatic atmosphere are computed using a cooling time argument. Briefly, we adopt a collisional ionization equilibrium cooling function computed for the metallicity of the hot halo gas (which in turn is computed self-consistently from the stellar yield and mass outflow rates from galaxies). Using this cooling function, we estimate the total energy radiated by an element of gas at each radius in the hot atmosphere. This radiated energy is therefore accumulated over the entire lifetime of a given gas element in the hot halo. This radiated energy is compared to the thermal energy of the gas. We define a cooling radius as the point where these two energies are equal and assume that all gas inside of that radius has been able to cool and accrete on to the central galaxy. The rate of growth of the cooling radius coupled with the assumed cored-isothermal density profile of the hot atmosphere then gives a mass accretion rate on to the galaxy.

When included, SN-driven outflows are modelled by assuming an outflow rate, $\dot{M}_{\rm out}$, proportional to the star formation rate, $\dot{M}_\star$. Specifically,

$$\dot{M}_{\rm out} = \beta \dot{M}_\star, \tag{5}$$

where $\beta = (V/V_{\rm hot})^{-\alpha_{\rm hot}}$, where $V_{\rm hot}$ and $\alpha_{\rm hot}$ are model parameters (for which we adopt $\alpha_{\rm hot} = 3.52$ and $V_{\rm hot} = 336\,{\rm km\,s^{-1}}$ for quiescent star formation and $V_{\rm hot} = 305$ for bursts of star formation) and $V$ is the characteristic circular velocity of the galaxy component (disc or spheroid) in question. The star formation rate itself is computed using

$$\dot{M}_\star = \epsilon_\star \frac{M_{\rm gas}}{\tau_{\rm disc}} \tag{6}$$

for discs, where $M_{\rm gas}$ is the mass of gas in the disc, $\tau_{\rm disc} = R_{\rm disc}/V_{\rm disc}$ is the disc dynamical time (with $R_{\rm disc}$ and $V_{\rm disc}$ being the half-mass radius of the disc and circular velocity at that radius, respectively) and $\epsilon_\star = 0.049$ is an efficiency parameter. In bursts a similar expression is used, but with a much larger efficiency ($\epsilon_\star = 0.5$), and the disc dynamical time is replaced with the corresponding dynamical time of the spheroidal component of the galaxy.

All of the models in the work incorporate feedback from active galactic nuclei (AGNs). In Benson & Bower (2010), AGN feedback works by shutting down cooling from a hydrostatic halo, if the energy output of the AGN exceeds the cooling luminosity of that halo. While this feedback channel is only operative in haloes, which are well above the critical mass-scale at which the transition from rapid to slow cooling occurs, a galaxy can accrete via the 'cold mode' even while experiencing active AGN feedback. As such, our new treatment of accretion shocks may allow for a higher gas accretion rate on to massive galaxies.

## 3 RESULTS

We find that values of $\epsilon_{\rm s,crit} = 0.0126$ (the value found by Birnboim & Dekel 2003), $\Delta\epsilon = 0.01$ and $\Gamma_{\rm rapid} = 2.5$ give results, which are in reasonable agreement with those found in numerical simulations. It should be noted that these correspond to a very simple model with little fine-tuning – $\epsilon_{\rm s,crit}$ is as predicted by a simple analytic argument, $\Gamma_{\rm rapid} = 2.5$ implies that the cold reservoir is depleted on a fraction of the free-fall time-scale as expected and $\Delta\epsilon = 0.01$ implies a rapid transition from hot to cold mode.

In particular, Fig. 1 shows a comparison with the results from hydrodynamical simulations as reported by Kereš et al. (2009). The fraction of halo gas presently in the cold reservoir is indicated by the black points as a function of halo mass and for four different redshifts. The blue lines indicate the median cold fraction at each halo mass, while the red lines indicate the corresponding hot fraction. For comparison, the magenta lines show the cold fraction as reported by Kereš et al. (2009). For reference, the green lines indicate the mass at which haloes transition from rapid- to slow-cooling regimes in the original GALFORM implementation. Clearly this occurs at a lower mass-scale than in our new implementation based on equation (1). Since most other semi-analytic models adopt a similar

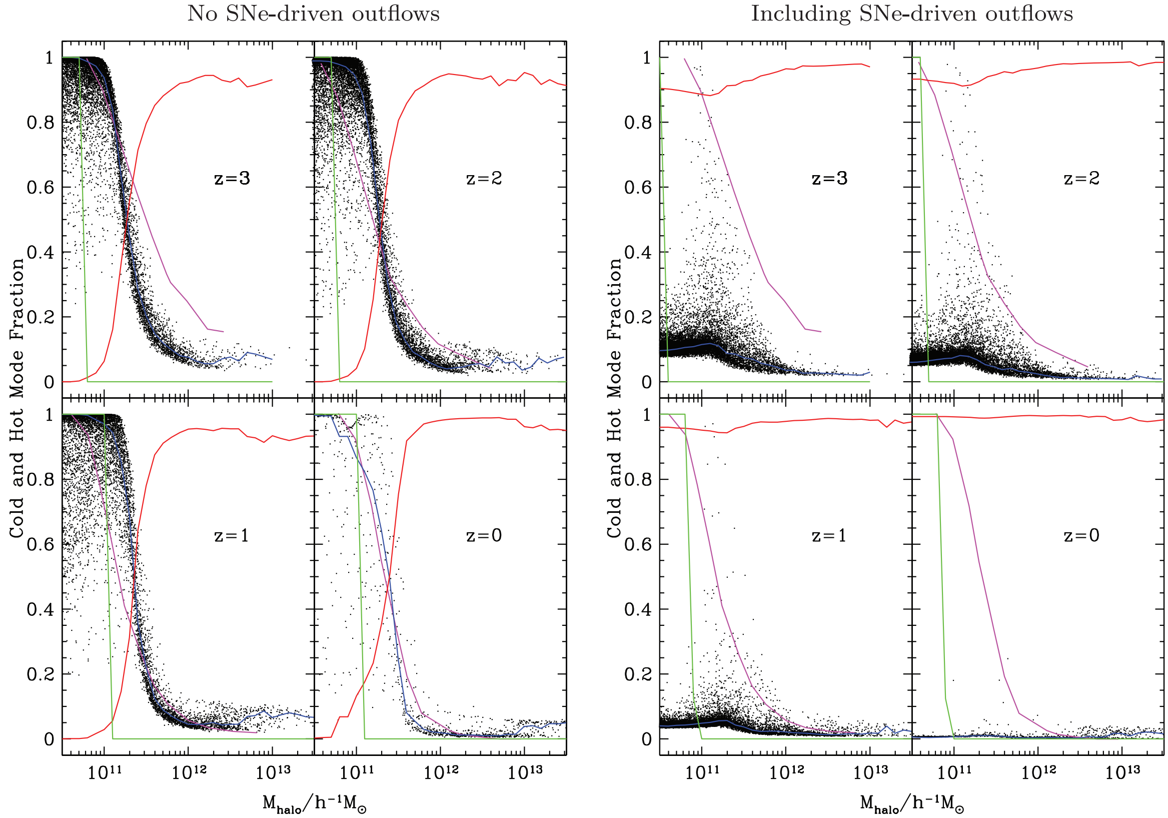


**Figure 1.** The fraction of diffuse (i.e. non-galactic) gas in haloes in the cold reservoir is shown by the black points as a function of halo mass and for four different redshifts as indicated in each panel. The solid blue line shows the median cold gas fraction at each halo mass. The red line shows the corresponding median hot gas fraction. The magenta lines show the equivalent cold gas fraction from the simulations of Kereš et al. (2009). The green lines indicate the transition from rapid to slow cooling in the original GALFORM implementation (haloes in the rapid-cooling regime are assigned a cold gas fraction of 1; those in the slow-cooling regime are assigned a value of 0). Left-hand panel: results when SN-driven outflows are not included. Right-hand panel: results when SN-driven outflows are included. The relative paucity of model points at $z = 0$ reflects the number of merger tree realizations, which we constructed – at higher redshifts, we include all progenitor haloes in these trees giving many more points.

criterion to that adopted in our original implementation, we expect their transition to occur at a similar mass-scale. The transition mass-scale evolves with redshift in almost the same way in the original and new implementations of the rapid/slow cooling transition. The differences are likely due to the fact, discussed in Section 2, that the original criterion depends upon the past history of the halo gas, while the new criterion depends only upon the instantaneous, pre-shock gas properties.

In the left-hand panel, we show results from our model with SN-driven outflows switched off. The simulations of Kereš et al. (2009) do not create strong outflows, which may be expected to significantly alter the relative fractions of cold and hot gas in haloes. Therefore, the fairest comparison to their results is made with outflows removed from our model. Our model produces results in reasonable agreement with those of Kereš et al. (2009), correctly finding the characteristic transition mass as a function of redshift and producing approximately the correct width of the transition at $z = 0$. Our model transitions somewhat too quickly at higher redshifts, which may be indicative of a break-down in the assumption of spherical symmetry in the model of Birnboim & Dekel (2003) and the ability of dense filaments to penetrate through the hot halo, an assumption which should be more valid at low redshifts. This may cause our model to somewhat underestimate the accretion rate via the cold channel at high redshifts. We will return to this point below. In the right-hand panel of Fig. 1, we switch SN-driven outflows back on in our model, under the assumption that gas driven out of galaxies will populate the hot component of the halo. SN feedback has no explicit effect on the shock stability calculation in our model and so does not affect the total mass of gas accreted into the halo without experiencing an accretion shock at the virial radius. However, we find that the cold fraction is greatly suppressed in this case, since the haloes now contain significant quantities of reheated gas, which dilutes the cold gas. This assumption could be incorrect, with outflowing gas condensing into small, dense clouds; our goal here is merely to indicate how important outflows *could* be in establishing the cold/hot accretion balance in haloes.

In Fig. 2, we explore the consequences of our new implementation of the rapid-cooling regime for the build-up of the galaxy population, compared with a model in which an accretion shock is assumed to always occur at the virial radius. The left-hand panel shows the volume-averaged star formation rate in models with and without cold-mode accretion (blue and red lines, respectively) in cases with and without SN-driven outflows (thick and thin lines, respectively). Of course, the models lacking SN-driven outflows produce far too high star formation rates as expected, but are shown for reference. Comparing the blue and red lines, the models have almost identical star formation rates at high redshifts, irrespective of the inclusion of our new treatment of accretion shocks both when SN-driven outflows are included and when they are not. At low redshifts, the models including our new treatment of accretion

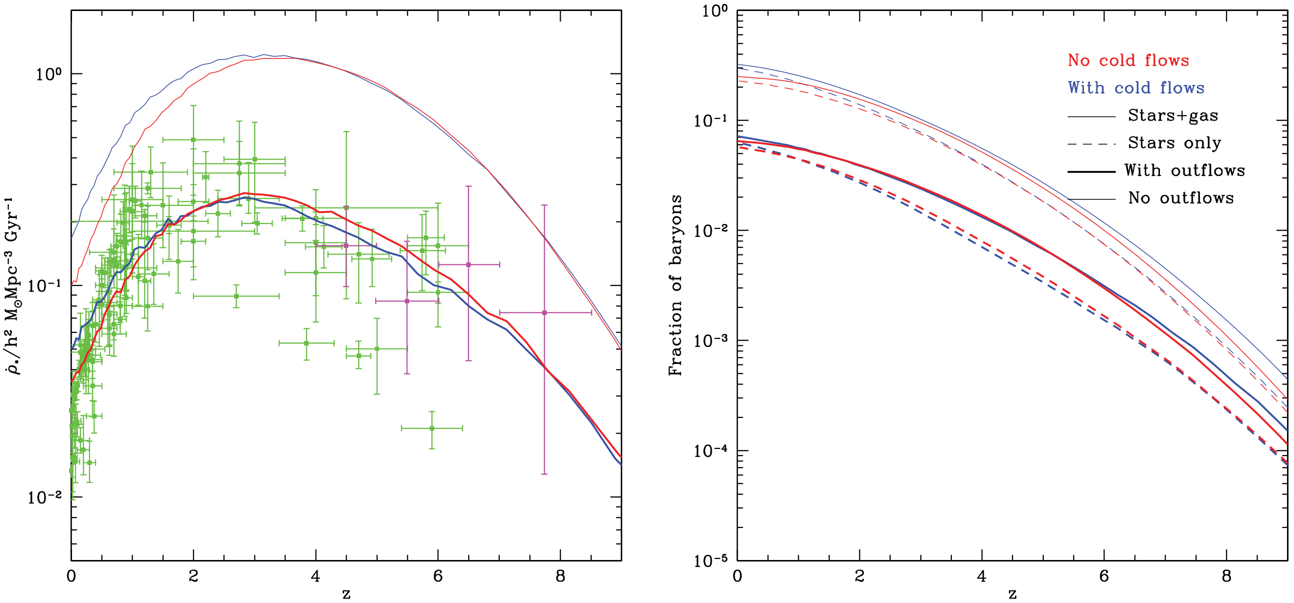


**Figure 2.** Left-hand panel: the volume-averaged star formation history as a function of redshift. Green points show observational estimates from a variety of sources as compiled by Hopkins (2004), while magenta points show the star formation rate inferred from gamma-ray bursts by Kistler et al. (2009). The red lines show results for models with the original treatment of the rapid-cooling regime, while the blue lines are for models including the rapid-cooling regime. The thick lines indicate models including SN-driven outflows, while the thin lines are for models without such outflows. Right-hand panel: the fraction of baryons locked up into stars (dashed lines) and stars plus galactic gas (solid lines) as a function of redshift. The thick and thin lines correspond to models with and without SN-driven outflows. The red lines show models without the rapid-cooling regime, while the blue lines correspond to models that include the rapid-cooling regime.

shocks have higher star formation rates. We note that the original GALFORM model, including the previous estimate of the rapid to slow cooling transition and including SN-driven outflows, has been constrained to fit a variety of observational data, including those shown in this figure. The addition of our new treatment of the rapid-cooling regime worsens the agreement with the data, but this should not be taken as evidence against this model – it is highly probable that adjustments in other parameters could restore good agreement in the new model. Comparing the thick and thin lines, the figure emphasizes the importance of using the correct galaxy formation physics to assess the importance of the rapid-cooling regime. The effects of stellar winds and outflows are far more important (but less certain) than the distinction between slow and rapid cooling.

The right-hand panel of Fig. 2 shows the fraction of baryons (averaged over the entire universe) in the form of stars (dashed lines) or stars and interstellar medium (ISM) gas (solid lines) as a function of redshift for the same set of four models. Considering first models with no SN-driven outflows, the fraction of baryons in stars is identical in models with and without the rapid-cooling regime at high redshifts as expected from the star formation rate. However, the fraction of baryons in stars plus ISM gas is actually higher in the model including the rapid-cooling regime. This occurs because the rapid-cooling regime is more efficient at getting gas into the galaxy phase but, in our model, results in galaxies with longer star formation time-scales due to a higher angular momentum content (and consequently longer disc dynamical times). As discussed in Section 2, this is because accretion from a hydrostatic halo always delivers the lowest angular momentum material to a galaxy first. Accretion via cold filaments is therefore able to get more angular momentum into a galaxy earlier. This is illustrated in Fig. 3, where we show the dynamical times of galactic discs as a function of the halo mass at $z = 6$. The dashed lines compare the effect of the rapid-cooling regime on dynamical time in the absence of SN-driven

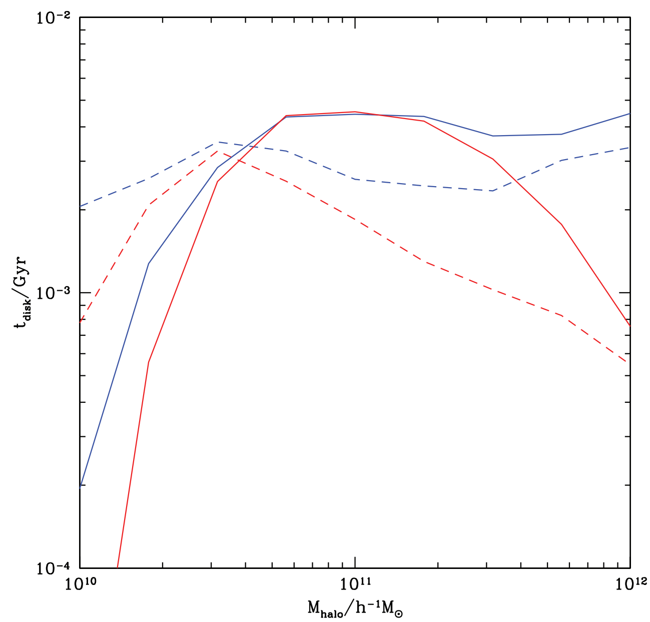

**Figure 3.** The median dynamical time of galactic discs as a function of the halo virial mass at $z = 6$. The dashed lines show models with no SN-driven outflows, while the solid lines include SN-driven outflows. The red lines show results for models with the old treatment of the rapid-cooling regime, while the blue lines are for models including the rapid-cooling regime.

outflows, while the solid lines are for models that include SN-driven outflows. In high-mass haloes, the rapid-cooling regime increases the disc dynamical time by a factor of 3 or more. In models that include SN-driven outflows, the effect is weaker and the difference in the total gas content is only significant at higher redshifts. By

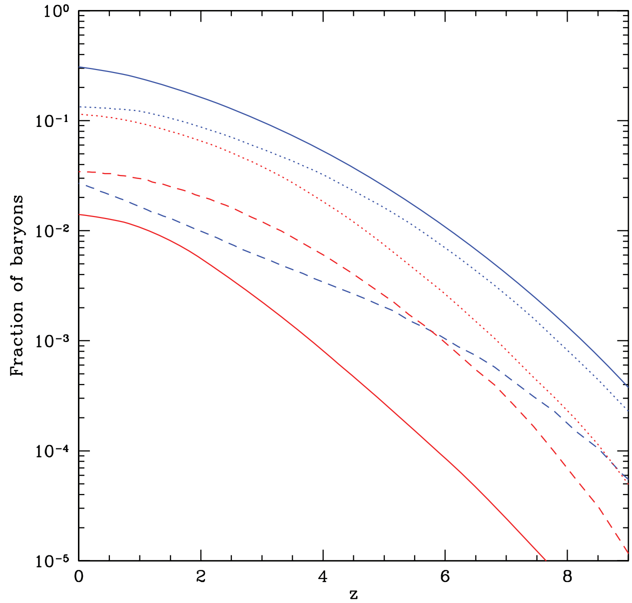


**Figure 4.** The fraction of baryons in galaxies (including both stars and gas) as a function of redshift. The blue lines show baryons that arrived 'cold' via the rapid-cooling regime (and which must have shocked close to the galaxy), while the red lines show baryons that arrived via cooling from a hot, hydrostatic halo. The solid lines show a model with no SN feedback, while the dashed lines show a model including such feedback. Note that in the model including feedback, gas, which first arrived in a galaxy via the 'cold' channel, can later be ejected by winds, incorporated into the hydrostatic halo, and later cools and flows back into the galaxy. We include such cases in the hot component (red line), since in our model, we are unable to distinguish gas in the hot halo, which originally accreted cold from that which arrived there directly from the IGM. For comparison, the dotted lines show the partitioning between rapid- and slow-cooling regimes in the original model, with no outflows included.

$z = 0$, in models with SN-driven outflows, the total mass of baryons in stars and ISM gas differs only slightly between models with and without a more detailed treatment of the rapid-cooling regime.

Fig. 4 shows directly the relative contributions of cold, filamentary infall and cooling from a hot, hydrostatic halo as a function of redshift. When SN feedback is ignored (solid lines), the cold infall (blue) line is the dominant channel for getting baryons into galaxies (when averaged over all galaxies), exceeding the slow-cooling channel by a factor of around 10. If SN feedback is included (dashed lines), the picture changes. We note that when we include feedback, gas which first arrived in a galaxy via the 'cold' channel can be ejected by winds, incorporated into the hot atmosphere halo and later cools and flows back into the galaxy. We include such gas in the hot component (red line), since in our model, we are unable to distinguish gas in the hot halo, which originally accreted cold from that which arrived there directly from the IGM. Thus, some material judged as being contributed by the hot component will have originally arrived via the cold channel. In any case, the two channels are seen to be much more comparable. This arises primarily because of the cycling of mass from the galaxy into the hot halo and back into the galaxy again, which moves mass from the cold to the hot channel in this figure, but also partially because outflows enrich the hot halo gas in metals, boosting their cooling rate. For comparison, we also show in Fig. 4 the division between rapid- and slow-cooling regimes in the original model in the case of no outflows. The total amount of mass cooled is comparable to that in our new calculation (solid lines), but the contribution from the slow regime (i.e. hot mode) is much greater than in the new calculation. This shows a dramatic difference in the nature by which most gas makes its way into galaxies.

Finally, Fig. 5 shows instantaneous accretion rates via the cold and hot channels on to central galaxies as a function of the galaxy's dark matter halo mass. The left-hand column shows a model with no SN-driven outflows and this can be compared, for example, to fig. 7 of Kereš et al. (2009). Our model can be seen to predict, qualitatively and quantitatively, similar accretion rates as found in hydrodynamical simulations, although it may overestimate the rate of cold-mode accretion in the most massive haloes by factors of a few. Additionally, as suspected from Fig. 2, our model slightly underestimates accretion rates via the cold channel at high redshifts, but, given the simplicity of the model, we consider this to be an acceptable level of agreement. One significant difference with the results of Kereš et al. (2009) is the lack of a significant contribution from the hot channel in haloes with masses above around $10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$. This is a consequence of the inclusion of AGN feedback in our model, which effectively shuts down the hot channel in such haloes. When outflows are included (right-hand column), the cold accretion rates are unaffected, but accretion rates from the hot channel are boosted significantly. This is due to the fact that outflows refill the hot gas reservoir, resulting in more gas available to cool, and enrich it in metals, boosting cooling rates. It should be noted that the feedback in Benson & Bower (2010) is particularly strong, a result of the need to produce a flat faint-end slope in the luminosity function. This leads to feedback rates that are difficult to justify on energetic grounds, if feedback is caused by SNe. Therefore, more energetically feasible feedback would likely result in lower rates of hot accretion.

## 4 DISCUSSION AND CONCLUSIONS

Semi-analytic models of galaxy formation have always made a distinction between rapid and slow regimes of cooling. With the recent attention on this aspect of galaxy formation, we have returned to the issue of how such models determine in which regime a given halo finds itself. In the rapid-cooling regime, we may expect gas accreted in the rapid-cooling regime to be unable to shock, until it reaches a radius much smaller than the virial radius of the halo, instead penetrating any hot atmosphere as a cold filament (or perhaps as clouds of dense gas associated with resolved subhaloes streaming along that filament). We have described a revised implementation of the role of accretion shocks in a semi-analytic model of galaxy formation, using a previously proposed analytical model with no modifications. Compared to our previous implementation, this new approach results in a shift in the transition mass-scale between rapid- and slow-cooling regimes (due to the new calibration of the numerical coefficients, which appears in the criterion for comparing cooling and free-fall time-scales) and a more rapid funneling of gas into galaxies in the rapid-cooling regime. Without any adjustment this model provides an excellent match to results from numerical simulations – this could no doubt be further improved with some fine-tuning of the model, but our aim here is to demonstrate that the simulation results can be encapsulated by an easy-to-implement way.

We find that the inclusion of SN-driven outflows has a dramatic effect on the fraction of cold-mode gas present in dark matter haloes – much more so than switching the revised treatment of accretion shocks on or off – suggesting that simulations must account for such outflows before they can make robust predictions for the effects of accretion in the 'cold accretion' regime. We caution that our

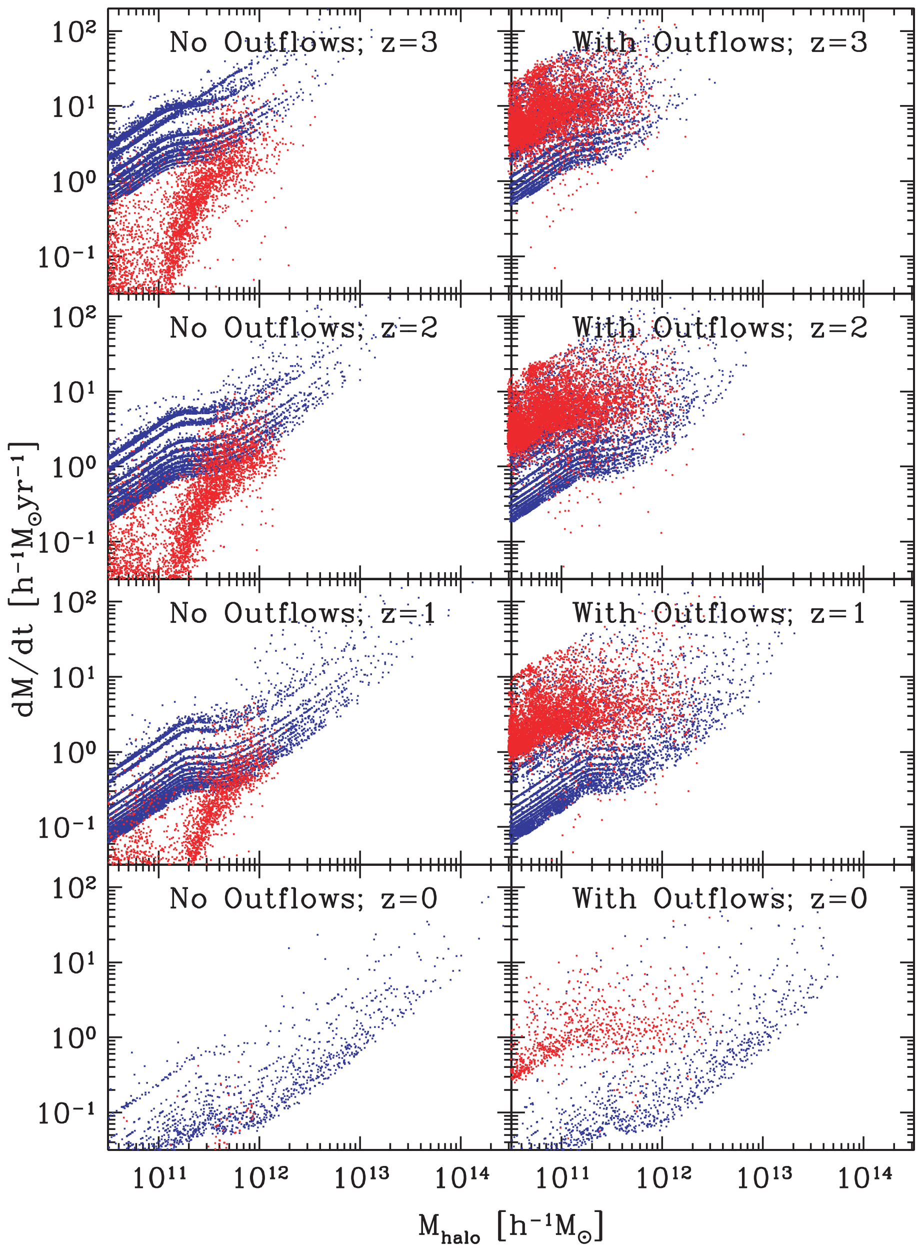


**Figure 5.** Accretion rates via cold (blue) and hot (red) channels as a function of the halo mass. The left-hand column shows results for a model with no outflows, while the right-hand column includes SN-driven outflows. Each row shows results for a different redshift, as labelled in each panel. Only central galaxies are shown. The quantization of accretion rates is a consequence of the finite resolution of the merger trees used in this coupled with the fixed time-steps over which accretion rates are computed. We have checked that reducing the size of time-steps and/or increasing the resolution of the merger trees does not affect our conclusions.

conclusions regarding the effect of SN feedback are specific to our particular implementation in which material in the outflow is deposited into the hot atmosphere of the halo after a time comparable to the dynamical time of the halo. Some observations (e.g. Steidel et al. 2010) suggest a different scenario in which the outflowing material is cold and moving with sufficient velocity that it may well escape the halo entirely. Numerical simulations also suggest other plausible feedback modes. For example, Brooks et al. (2009) include SN-driven outflows in their hydrodynamical simulations, which they show affects the properties of low-mass galaxies, but nevertheless find that ‘cold’ accretion remains the dominant mode in high-redshift galaxies. Similarly, Oppenheimer et al. (2009) incorporate SN-driven feedback into their simulations and demonstrate that this can produce a reasonable galaxy mass function in the sub-Milky Way mass regime (at $z \gtrsim 2$), but that this does not affect the rate of accretion via cold filaments. Understanding the interaction of feedback and filamentary accretion is clearly crucial to developing insight into how galaxies are fueled.

The inclusion of an updated treatment of the rapid-cooling regime makes little difference to luminosity functions and galaxy sizes at $z = 0$, implying that results from earlier semi-analytic models are still valid. Without SN-driven outflows, the new treatment results in a significant increase in the mass and luminosity of brighter galaxies, which are able to gain some mass through the rapid-cooling regime even when their slow cooling fuel supply has been effectively shut down by feedback from AGNs. The rapid-cooling regime is found to be effective at getting gas into the galaxy phase at high redshifts, potentially allowing for high rates of star formation during these epochs. However, we find that the star formation rate at $z \gtrsim 3$ is largely unaffected, when we included a treatment of the rapid-cooling regime, due to the fact that the galaxies that form are of larger and lower density than they would be, if the rapid-cooling regime were neglected. We caution that the treatment of angular momentum delivery to galaxies via filaments of cold gas in the rapid-cooling regime is therefore of crucial importance to assessing its impact on star formation rates. To date, this has not been studied in detail in hydrodynamical simulations, but should be in order to refine our understanding of how the rapid-cooling regime affects galaxy formation.

In summary, while it is desirable to include the rapid-cooling regime in semi-analytic models using the latest calibrations of the transition mass-scale, current and previous semi-analytic models treat the division accurately enough for current applications – the accuracy of these models (and of hydrodynamical simulations) is much more limited by their treatment of feedback and frequent assumption that baryon content of haloes is fixed (Bower, McCarthy & Benson 2008; Crain et al. 2009). Not surprisingly, given that this distinction has always been present in such models, our new implementation does not qualitatively change our picture of galaxy formation, at least at the coarse-grained level studied here. More detailed results (luminosity function shapes, distribution of sizes, and formation times) will require a significantly more detailed implementation of the rapid-cooling regime, including how it delivers angular momentum to galaxies (e.g. Navarro et al. 2009). Further calibration from numerical simulations should improve the accuracy of results. However, we have seen that the inclusion of updated prescriptions for the rapid-cooling regime into our semi-analytic model does not produce any surprises. The changes in star formation rates and stellar mass fractions, particularly in the low-redshift universe, are relatively small, smaller than other uncertainties in the galaxy formation theory. In particular, they are much smaller than the differences in galaxy properties created by changes to the feedback effects of galactic winds and AGNs.

## ACKNOWLEDGMENTS

AJB acknowledges the support of the Gordon & Betty Moore Foundation and thanks Alyson Brooks, Brant Robertson, Yuval Birnboim and Simon White for numerous helpful conversations. We thank Carlos Frenk, Shaun Cole, Cedric Lacey and Carlton Baugh for allowing us to make use of the GALFORM semi-analytic model for this work.

## REFERENCES

Agertz O. et al., 2007, MNRAS, 380, 963
Benson A. J., Bower R. G. 2010, MNRAS, 405, 1573
Benson A. J., Bower R. G., Frenk C. S., White S. D. M., 2000, MNRAS, 314, 557
Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, MNRAS, 320, 261
Bertschinger E., 1985, ApJS, 58, 39
Binney J., 1977, ApJ, 215, 483
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Book L. G., Benson A. J., 2010, ApJ, 716, 810
Bower R. G., McCarthy I. G., Benson A. J., 2008, MNRAS, 390, 1399
Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396
Cattaneo A., Dekel A., Devriendt J., Guiderdoni B., Blaizot J., 2006, MNRAS, 370, 1651
Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Crain R. A. et al., 2009, MNRAS, 399, 1773
Crain R. A., McCarthy I. G., Frenk C. S., Theuns T., Schaye J., 2010, MNRAS, 407, 1403
Croton D. J. et al., 2006, MNRAS, 365, 11
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A. et al., 2009, Nat, 457, 451
Fardal M. A., Katz N., Gardner J. P., Hernquist L., Weinberg D. H., Davé R., 2001, ApJ, 562, 605
Forcada-Miro M. I., White S. D. M., 1997, preprint (astro-ph/9712204)
Goerdt T., Dekel A., Sternberg A., Ceverino D., Teyssier R., Primack J. R., 2010, MNRAS, 407, 613
Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., Pearce F. R., 2003, MNRAS, 338, 913
Hopkins A. M., 2004, ApJ, 615, 209
Katz N., Gunn J. E., 1991, ApJ, 377, 365
Katz N., Keres D., Dave R., Weinberg D. H., 2003, in Rosenberg J. L., Putman M. E., eds, The IGM/Galaxy Connection: The Distribution of Baryons at $z = 0$. Kluwer, Dordrecht, p. 185
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160
Kistler M. D., Yuksel H., Beacom J. F., Hopkins A. M., Wyithe J. S. B., 2009, ApJ, 705, 104
McCarthy I. G. et al., 2007, MNRAS, 376, 497
Navarro J. F., Sales L. V., Schaye J., Vecchia C. D., Springel V., 2009, in Jogee S., Marinova I., Hao L., Blanc G. A., eds, ASP Conf. Ser. Vol. 419, Galaxy Evolution: Emerging Insights and Future Challenges. Astron. Soc. Pac., San Francisco, p. 10
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2009, MNRAS, 406, 2325
Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557
Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541
Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N. A., Bogosavljevi M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289
Tozzi P., Norman C., 2001, ApJ, 546, 63
Voit G. M., Balogh M. L., Bower R. G., Lacey C. G., Bryan G. L., 2003, ApJ, 593, 272
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, MNRAS, 335, 762

This paper has been typeset from a TEX/LATEX file prepared by the author.